AD-A059 924   NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA   F/G 12/1

RAPID EQUALIZATION OF HIGHLY DISPERSIVE CHANNELS USING ADAPTIVE--ETC(U)

APR 78   E H SATORIUS, S T ALEXANDER

UNCLASSIFIED   NOSC/TR-249   NL
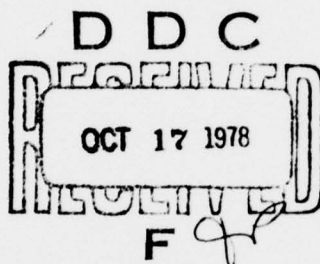
1 OF 1
AD
A059924

END
DATE
FILMED
12-78
DDC

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL

NOSC

Technical Report 249

RAPID EQUALIZATION OF HIGHLY DISPERSIVE CHANNELS USING ADAPTIVE LATTICE ALGORITHMS

EH Satorius
ST Alexander
30 April 1978

**NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152**

78 09 28 055

NOSC

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

A N ACTIVITY OF THE NAVAL MATERIAL COMMAND

RR GAVAZZI, CAPT, USN
Commander

HL BLOOD
Technical Director

ADMINISTRATIVE INFORMATION

Released by
R H HEARN, Head
Electronics Division

Under Authority of
D A KUNZ, Head
Fleet Engineering Department

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NOSC/TR-249	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* RAPID EQUALIZATION OF HIGHLY DISPERSIVE CHANNELS USING ADAPTIVE LATTICE ALGORITHMS		5. TYPE OF REPORT & PERIOD COVERED Research January to March 1978
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) E.H. Satorius S.T. Alexander		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Ocean Systems Center San Diego, CA 92152		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62611N F11101 XF11101100
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Electronic Systems Command Washington, D.C.		12. REPORT DATE 30 April 1978
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Adaptive lattice algorithm
Channel equalization
Orthogonalization properties

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This report presents a study of adaptive lattice algorithms as applied to channel equalization. The orthogonalization properties of the lattice algorithms make them promising for equalizing channels with heavy amplitude and/or phase distortion. Furthermore, unlike the majority of other orthogonalization algorithms, the number of operations per update for the adaptive lattice equalizers is linear with respect to the number of equalizer taps.

DD $_{1\ \text{JAN}\ 73}^{\text{FORM}}$ 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

CONTENTS

1

INTRODUCTION

The problem of equalizing a channel whose channel-correlation matrix has a large eigenvalue spread is well-known. Adaptive gradient algorithms[1,2] are among the simplest to implement, but the rate of convergence (ROC) of these algorithms is determined largely by the ratio of the maximum to minimum eigenvalues of the channel-correlation matrix.[3] Alternative algorithms have been proposed which orthogonalize the above matrix. In particular, Godard[4] through application of Kalman filter theory, has derived an adaptive self-orthogonalizing algorithm which has extremely rapid convergence properties. As discussed in Reference 5, the Godard algorithm involves estimating the inverse of the channel-correlation matrix through an iterative matrix equation. Even though the Godard algorithm converges rapidly, the number of operations per update for this algorithm depends on the square of the number of equalizer taps, which creates implementation difficulties for large length equalizers. Gitlin and Magee[5] have proposed an adaptive self-orthogonalizing algorithm which provides a compromise between computational complexity and speed of convergence. This algorithm consists of approximating the inverse of the channel-correlation matrix by a Toeplitz matrix and involves only one matrix multiplication. Reference 5 presents a comparison between some of the different adaptive orthogonalizing algorithms.

A relatively new class of adaptive algorithms also provides self-orthogonalizing capabilities and only requires a number of operations per update that depends linearly on the length of the filter.[6,7,8,9,10,11] These algorithms are called adaptive lattice (AL) algorithms and generate a set of orthogonal signal components which can be used as inputs to equalizer gain controls. These components are generated through a Gram-Schmidt type of orthogonalization.[9,10,11] These AL algorithms have been proposed for use in such areas as speech, sonar signal processing, noise cancelling and parameter estimation,[6,7,8,9,10,11] and Makhoul in References 10 and 11 has suggested they be used also in adaptive equalization. This report examines, via computer simulation, the performance of AL algorithms as applied to channel equalization.

ADAPTIVE LATTICE EQUALIZATION

As pointed out in Reference 5, the main component in the majority of adaptive equalization algorithms is the estimated gradient tap adjustment algorithm. This is true also for the AL algorithms examined in this report and, therefore, we present a brief review of the estimated gradient algorithm.

ESTIMATED GRADIENT ALGORITHM

The estimated gradient tap adjustment algorithm is given by:

$$\underline{C}_{n+1} = \underline{C}_n - \alpha_n e_n \underline{X}_n \ . \tag{1}$$

In Equation 1, $\underline{C}_n$ is the N-length vector representing the n-th iteration estimate of the optimum (minimum mean square error) N-tap equalizer; α_n is a positive parameter denoted as the step size; e_n is the instantaneous difference between the equalizer output (y_n) and the value of the transmitted data symbol (a_n); and $\underline{X}_n$ is the N-length vector of received data

samples in the equalizer delay line at the n-th iteration.* The components of X_n will be denoted as $[X_n]_k = x_{n-k+1}$ where $k = 1, \ldots, N$. The received data samples x_n are given as the superposition of a corruptive white noise sequence w_n and the output of a linear channel filter h_i operating upon the transmitted data symbols:

$$x_n = \sum_i a_{n-i} h_i + w_n \quad . \tag{2}$$

In Equation 2, the h_i represent the channel impulse response and the noise sequence w_n is defined to have a variance σ^2.

The convergence properties of Equation 1 have been studied by a number of authors[2,3,12] under different assumptions concerning the dependence of α_n upon n. In many applications of Equation 1 to digital adaptive equalization, α_n is held constant[3] and Equation 1 becomes the familiar least-mean-squares (LMS) algorithm.[2] As $n \to \infty$, the mean weight vector converges to the optimal tap vector given by

$$\lim_{n \to \infty} E(C_n) = C_{opt} \quad , \tag{3}$$

where $E(\cdot)$ denotes expectation and C_{opt} is the optimum weight vector given by the discrete Wiener matrix equation

$$C_{opt} = A^{-1} B \quad . \tag{4}$$

In Equation 4, A is the $N \times N$ positive-definite Toeplitz channel-correlation matrix

$$A = E(X_n X_n^T) \tag{5}$$

and B is the N-length vector

$$B = E(a_n X_n) \quad . \tag{6}$$

As the tap vector evolves according to Equation 1, the mean square error $\epsilon(n)$ evolves as

$$\epsilon(n) = \epsilon_{opt} + E((C_{opt} - C_n)^T A (C_{opt} - C_n)) \tag{7}$$

where ϵ_{opt} is the minimum mean square error (MMSE) and is given by

$$\epsilon_{opt} = 1 - B^T C_{opt} \quad . \tag{8}$$

It is assumed the data symbol sequence is uncorrelated and has unity power.

As discussed in References 3 and 5, the convergence of ϵ_n is strongly dependent upon the ratio R of the largest-to-smallest eigenvalues of the A matrix; i.e., $R = \lambda_{max}/\lambda_{min}$. Therefore, large values of R (heavy channel distortion) can lead to excessively long convergence times when Equation 1 is used to update the C_n. Basically this is because the components of X_n generally are not orthogonal. This implies that premultiplying the estimated gradient $(e_n X_n)$ in Equation 1 by A^{-1} (or an estimate of A^{-1}) potentially will

---

*In this report we will consider only the training part of equalization when the transmitted data symbols are known at the receiver.

4

offer a significant time improvment in the convergence of Equation 1. This approach is employed in the algorithms of Godard[4] and Gitlin and Magee[5] and Chang[13] has suggested transforming $\underline{X}_n$ to a new vector $\underline{Z}_n$ with orthonormal components. Unfortunately, these algorithms have computational and/or storage requirements which grow as the square of the number of equalizer taps.

ADAPTIVE LATTICE ALGORITHMS

As an alternative to the self-orthogonalization algorithms discussed previously, Makhoul[10,11] has suggested the use of AL algorithms for adaptive equalization. Given an equalizer input sequence x_n, the AL algorithms generate an orthogonal set of signals which will be denoted as $b_m(n)$, where $m = 1, \ldots, N$. Although a number of AL algorithms have been proposed[10,11] for performing this orthogonalization, concentration will be directed to the basic lattice structure shown in Figures 1 and 2. This particular lattice structure was originally proposed by Itakura and Saito[15] for performing speech analysis. The orthogonalization of $\underline{X}_n$ is done in this lattice through the recursions:

$$b_1(n) = f_1(n) = x_n \tag{9a}$$

$$f_{m+1}(n) = f_m(n) - K_m \, b_m(n-1) \tag{9b}$$

$$b_{m+1}(n) = -K_m \, f_m(n) + b_m(n-1) \tag{9c}$$

where $m = 1, \ldots, N-1$. The $f_m(n)$ and $b_m(n)$ in Equation 9 are called the forward and backward error residuals of the lattice[10,11,14], respectively, and their properties will be discussed presently. The K_m in Equation 9 are known as the reflection coefficients and may be determined by a number of methods[14] which produce identical results when x_n is a statistically stationary sequence. The method used for choosing the K_m in this report was originally proposed by Burg[16] and consists of minimizing the sum of the variances of the backward and forward residuals, denoted as $E(f_{m+1}^2(n)) + E(b_{m+1}^2(n))$, with respect to K_m. The result for the K_m is given by[14]:

$$K_m = \frac{2\,E(f_m(n)\,b_m(m-1))}{E(f_m^2(n)) + E(b_m^2(n-1))} \quad, \quad 1 \leqslant m \leqslant N-1 \tag{10}$$

As discussed in References 6 through 11 and 14 and 15, when x_n is stationary the residuals $b_{m+1}(n)$ and $f_{m+1}(n)$ in Equation 9 are equivalent to the backward and forward error residuals of an m-point 1-step linear prediction filter. That is,

$$b_{m+1}(n) = x_{n-m} - \sum_{j=1}^{m} W_j^{(m)} \, x_{n-m+j} \quad, \quad 1 \leqslant m \leqslant N-1 \tag{11a}$$

$$f_{m+1}(n) = x_n - \sum_{j=1}^{m} W_j^{(m)} \, x_{n-j} \quad, \quad 1 \leqslant m \leqslant N-1 \tag{11b}$$
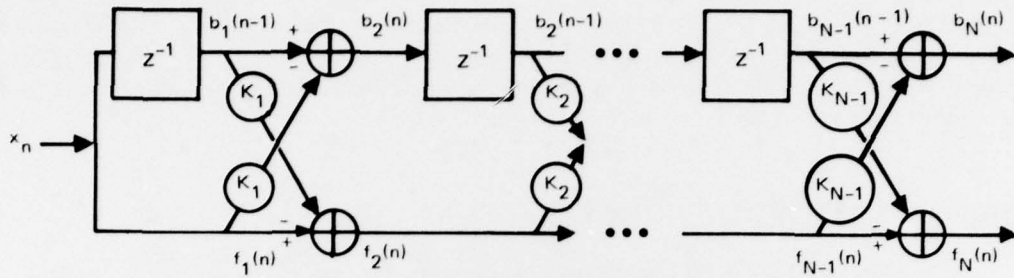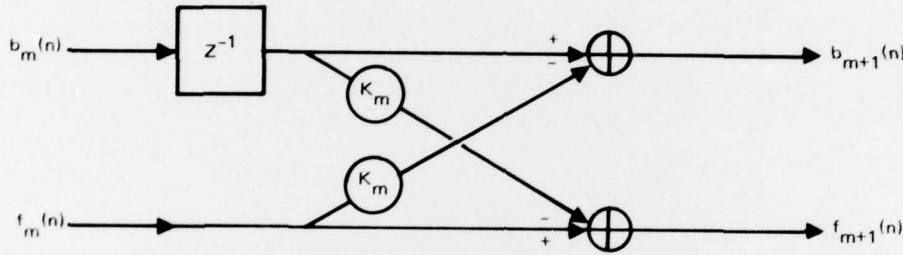
Figure 1. Basic lattice structure.

Figure 2. The m-th stage of the basic lattice structure.

where the $W_j^{(m)}$ are the 1-step predictor coefficients obtained from the normal equations[17]:

$$\sum_{j=1}^{m} W_j^{(m)} \phi(p - j) = \phi(p) \quad , 1 \leqslant p \leqslant m \quad . \tag{12}$$

In Equation 12, $A_{pj} = \phi(p - j)$ is the p, j-th element of the channel correlation matrix. Given the basic orthogonality property of MMSE residuals, namely,

$$E(x_{n-j} \, b_{m+1}(n)) = 0 \quad , \quad 0 \leqslant j \leqslant m - 1 \quad , \tag{13}$$

it is easily seen from Equation 11a that

$$E(b_m(n) \, b_j(n)) = \begin{cases} 0 \, , j \neq m \\ \epsilon_m \, , j = m = 2,3, \ldots, N \\ E(x_n^2) \, , j = m = 1 \end{cases} \tag{14}$$

where ϵ_{m+1} represents the MMSE of an m-point 1-step linear prediction filter. Therefore, the bm(n) represent a set of N orthogonal signals which can now be used as the inputs to equalizer gain controls.* However, it remains to be seen how the lattice algorithm of Equation 9 and Equation 10 can be implemented adaptively.

---

*It should be noted that an additional set of N orthogonal signals, $f_m(n + m - N)$, where $m = 1, \ldots, N - 1$, is also available from the lattice algorithm, as can be seen by using a similar argument as that which led to Equation 14.

IMPLEMENTATION

We will use the algorithm below, proposed by Makhoul,[10,11] which is a modification of the AL algorithm presented by Griffiths.[8] Specifically, K_m in Equations 9 and 10 is replaced by $K_m(n)$ and updated according to the following adaptive algorithm:

$$K_m(n + 1) = K_m(n) + \frac{\alpha}{\sigma_m^2(n)} \left\{ f_{m+1}(n) b_m(n-1) + f_m(n) b_{m+1}(n) \right\} ,$$

$$1 \leq m \leq N - 1 . \tag{15}$$

In Equation 15, α is the normalized step size of the adaptive algorithm and is restricted to $0 < \alpha < 2$ for stability[10,11]. Also, $\sigma_m^2(n)$ is the n-th iteration estimate of the sum $E(f_m^2(n)) + E(b_m^2(n-1))$. The $\sigma_m^2(n)$ are updated as follows:

$$\sigma_m^2(n) = (1 - \alpha) \sigma_m^2(n - 1) + \alpha \left\{ f_m^2(n) + b_m^2(n - 1) \right\} ,$$

$$1 \leq m \leq N - 1 . \tag{16}$$

Equations 9, 15, and 16 are the AL algorithm which provides the orthogonal signal components $b_m(n)$. Given the backward and forward residuals at the n-th update, $K_m(n+1)$ is computed from Equations 15 and 16. The backward and forward residuals at the (n+1)-st update are then computed from Equation 9 using the updated reflection coefficients $K_m(n+1)$ and so on.

Equations 15 and 16 provide noisy estimates of the optimal K_m values as given by Equation 10. The variance of these estimates is reduced as $\alpha \to 0$ (which also results in an increase in the convergence time of Equations 15 and 16). However, due to the successive orthogonalization which is intrinsic to the lattice structure, it is expected that the convergence rate of Equations 15 and 16 will not be limited by the ratio R, as is the case with Equation 1. This is indeed displayed by simulations presented in the next section, Stimulation Results.

Let us examine two specific adaptive algorithms for estimating the MMSE equalizer tap coefficients. The first algorithm, illustrated in Figure 3, has recently been proposed by Griffiths[9] and is given by:

$$V_1(n) = a_n - G_1(n) b_1(n) , \tag{17a}$$

$$V_m(n) = V_{m-1}(n) - G_m(n) b_m(n) , \quad 2 \leq m \leq N . \tag{17b}$$

The tap coefficients, $G_m(n)$, are updated according to:

$$G_m(n + 1) = G_m(n) + \frac{\alpha}{\gamma_m^2(n)} V_m(n) b_m(n) , \quad 1 \leq m \leq N , \tag{18a}$$

$$\gamma_m^2(n) = (1 - \alpha) \gamma_m^2(n - 1) + \alpha b_m^2(n) , \quad 1 \leq m \leq N . \tag{18b}$$
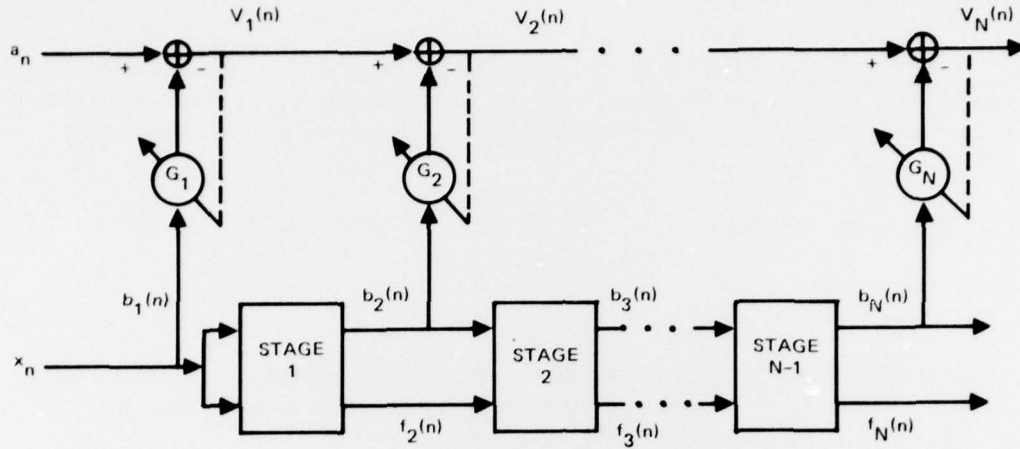
Figure 3. Adaptive lattice algorithm of Griffiths.[9]

As discussed in Reference 9, the sequence $V_m(n)$ represents the output error sequence of an m-tap equalizer. Therefore,

$$V_m(n) = a_n - y_n^{(m)} \quad , \tag{19}$$

where $y_n^{(m)}$ is the output sequence of the m-tap equalizer. Note that the value $\gamma_m^2(n)$ is the n-th iteration estimate of the power, $E(b_m^2(n))$, of the orthogonal signal component $b_m(n)$. This component is fed directly into the m-th equalizer tap, G_m, as shown in Figure 3. Griffiths[9] points out that the convergence rate of the overall AL algorithm (represented by Equations 9 and 15 through 18) should be relatively insensitive to the eigenvalue ratio R, especially when compared to Equation 1. This observation is valid, evidenced by the results of the simulations in the next section.

A second adaptive algorithm for adjusting the equalizer taps, suggested by Makhoul,[10,11] is illustrated in Figure 4. This algorithm is given by:

$$V_N(n) = a_n - \sum_{m=1}^{N} G_m(n) b_m(n) \quad , \tag{20}$$

where the tap coefficients are updated according to

$$G_m(n+1) = G_m(n) + \frac{\alpha}{\gamma_m^2(n)} [V_N(n) b_m(n)] \quad , \tag{21a}$$

$$\gamma_m^2(n) = (1 - \alpha)\gamma_m^2(n-1) + \alpha b_m^2(n-1) \quad , \tag{21b}$$

$$1 \leqslant m \leqslant N \quad .$$

As can be seen from Equations 20 and 21 and Figure 4, $V_N(n)$ represents the error output sequence from an N-tap equalizer.
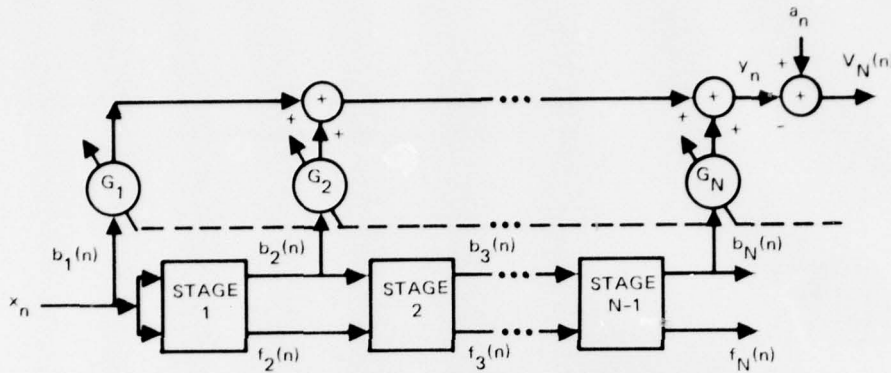
8

Figure 4. Adaptive lattice algorithm of Makhoul,[10,11]

The AL algorithm represented by Equations 9 and 15 through 18, will be referred to as AL1 and the AL algorithm represented by Equations 9, 15 through 16, 20 and 21, will be referred to as AL2. The main difference between the two algorithms is that in AL1 each individual error sequence ($V_m(n)$ for m=1, ...,N) is available, whereas AL2 only provides the N-tap equalizer output error sequence. This property of AL1, as noted in Reference 9, makes it potentially useful for purposes of determining the optimum number of taps for use in a time-varying environment. Specifically, since the time constant of the overall adaptive lattice configuration is proportional to the number of stages,[8] later stages will have larger time constants and, therefore, will not be able to track a highly dynamic input. Thus, the sequence of error expectations, $E\{V_m^2(n)\}$ for m=1, ...,N, will have a minimum for some m.

A number of other properties of adaptive and fixed-structure lattice algorithms are presented in References 8 through 11.

SIMULATION RESULTS

Results of AL1 and AL2 computer simulations are presented in this section for two channels representing heavy distortion (R = 11,21). In all simulations 11-tap equalizers were used and the symbol sequence a_n was a random sequence of bipolar signals ($a_n = \pm 1$), suitably delayed so that the optimal taps (given by Equation 4) were symmetric about the center of the equalizer. Also, the channel impulse response in all cases was the raised-cosine pulse, defined by:

$$h_i = \begin{cases} \frac{1}{2}(1 + \cos\{2\pi(i - N_o - 1)/W\}), & 1 \leq i \leq 2N_o + 1 & (22a) \\ \\ 0, \text{ otherwise} & (22b) \end{cases}$$

where W in Equation 22a was varied to provide different values for the eigenvalue ratio R.

The results of the simulations for the two different channels are presented in Figures 5 and 6. All plots were generated by ensemble averaging the squared error output of the equalizer over 200 individual learning curves. For purposes of comparison, the gradient algorithm of Equation 1 was also simulated. For the gradient algorithm, the step size was
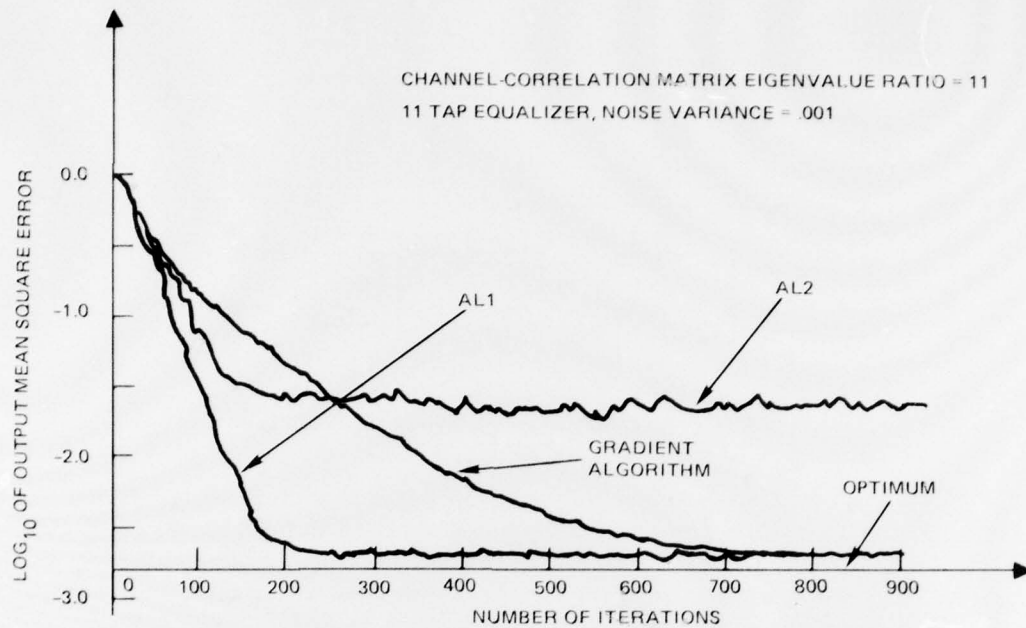
Figure 5. Comparison by simulation of convergence properties for eigenvalue ratio = 11.
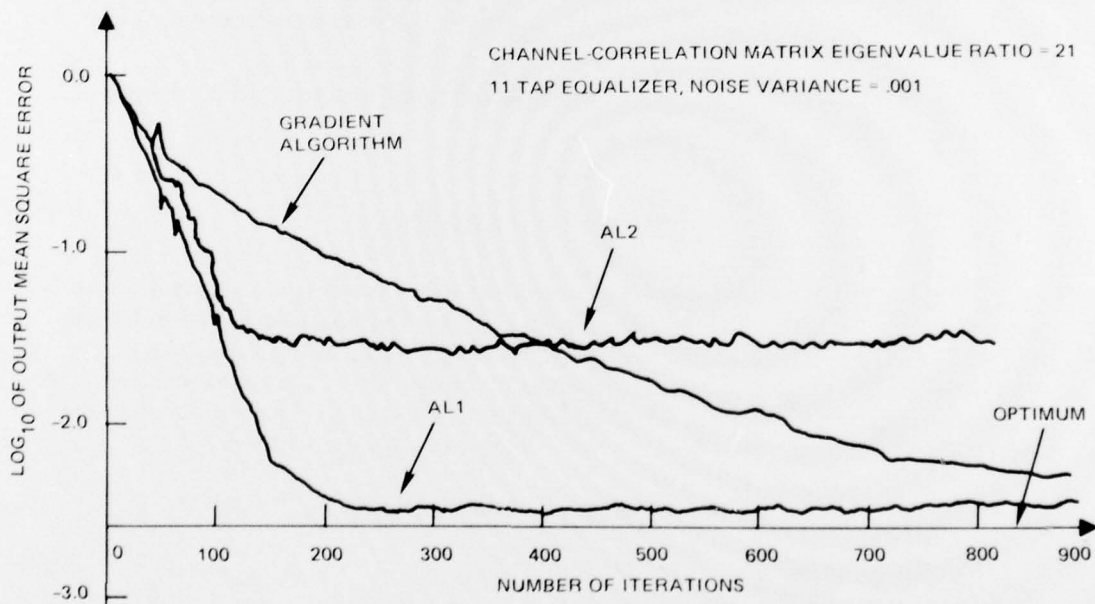

Figure 6. Comparison by simulation of convergence properties for eigenvalue ratio = 21.

chosen to be $\alpha = 0.02$; for the lattice algorithms α was chosen to be 0.025. Experimentation has shown that these values of α provided a good trade-off between quick convergence and stability. The initial equalizer tap values for each simulation was the zero vector. In addition, the initial values for the $K_m(n)$ coefficients in Equation 15 were zero, and the $\sigma_m^2(n)$ and the $\gamma_m^2(n)$ were initialized to unity.

The simulation results for AL channel equalization of high distortion communication channels are shown in Figures 5 and 6. Several points are of immediate interest. First, the initial convergence behaviors AL1 and AL2 are faster than for the gradient estimation equalizer, and are linear, rather than convex. Second, there is a noticeable difference in the steady-state MSE levels between AL1 (which updates the $G_j(n)$ according to the stage-wise errors $V_j(n)$) and AL2 (which updates the $G_j(n)$ using only the final error $V_N(n)$). These results do not necessarily establish the superiority of AL1 over AL2 for equalizer implementations since very little analysis has been done on optimizing the lattice and equalizer parameters. However, several different α values were used in addition to the one used for Figures 5 and 6 and a similar reduced MSE value resulted from AL1. Another important result is seen from comparing the AL1 curves in Figures 5 and 6. The AL1 equalizer converges for both the R=11 and R=21 channels in approximately equal times and thus exhibits the eigenvalue insensitivity which theory suggests. The gradient estimate equalizer, on the other hand, required an increasing number of iterations to reach convergence as the eigenvalue ratio was increased.

CONCLUSIONS

In this report, the application of adaptive lattice algorithms to channel equalization has been considered. Unlike the majority of proposed self-orthogonalizing algorithms, the AL algorithms only require a number of operations per update which is linear with respect to the number of equalizer taps. Furthermore, the rate of convergence of AL algorithms appears highly insensitive to the eigenvalue disparity of the channel-correlation matrix. Of the two AL algorithms investigated, the algorithm AL1 which minimized the stage-wise errors of the equalizer possesses a much lower steady-state MSE than the algorithm AL2 which minimized only the final equalizer error. This excess mean square error in steady state may be attributed to the fact that the equalizer tap coefficients may be noisier when minimizing only the final error. However, at present this relationship is not well understood and is an area for further investigation.

REFERENCES

1. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., Principles of Data Communication, New York: McGraw-Hill, 1968.

2. B. Widrow, "Adaptive Filters," in Aspects of Network and System Theory, R. Kalman and N. De Claris, Eds., New York: Holt, Rhinehart and Winston, pp. 563–587, 1971.

3. R. D. Gitlin, J. E. Mazo and M. G. Taylor, "On the Design of Gradient Algorithms for Digitally Implemented Adjustment Filters," IEEE Transactions on Circuit Theory, Vol. CT-20, pp. 125–136, March 1973.

4. D. Godard, "Channel Equalization Using a Kalman Filter for Fast Data Transmission," IBM Journal of Research and Development, pp. 267–273, May 1974.

5. R. D. Gitlin, F. R. Magee, "Self-Orthogonalizing Adaptive Equalization Algorithms," IEEE Transactions on Communications, Vol. COM-25, No. 7, July 1977.

6. M. D. Srinath and M. M. Viswanathan, "Sequential Algorithm for Identification of Parameters of an Autoregressive Process," IEEE Transactions on Automatic Control, pp. 542–546, August 1975.

7. R. Viswanathan and J. Makhoul, "Sequential Lattice Methods for Stable Linear Prediction," EASCON '76, Washington, D.C., paper 155, 1976.

8. L. J. Griffiths, "A Continuously-Adaptive Filter Implemented as a Lattice Structure," IEEE International Conference on Acoustics, Speech and Signal Processing, Hartford, CT, pp. 683–686, May 1977.

9. L. J. Griffiths, "An Adaptive Lattice Structure for Noise-Cancelling Applications," presented at the IEEE International Conference on Acoustics, Speech and Signal Processing, Tulsa, OK, April 1978.

10. J. Makhoul and R. Viswanathan, "Adaptive Lattice Methods for Linear Prediction," presented at the IEEE International Conference on Acoustics, Speech and Signal Processing, Tulsa, OK, April 1978.

11. J. Makhoul, "A Class of All-Zero Lattice Digital Filters: Properties and Applications," submitted to IEEE Transactions on Acoustics, Speech and Signal Processing, September 1977.

12. G. Ungerboeck, "Theory on the Speed of Convergence in Adaptive Equalizers for Digital Communications," IBM Journal of Research and Development, pp. 546–555, November, 1972.

13. R. W. Chang, "A New Equalizer Structure for Fast Start-Up Digital Communication," Bell System Technical Journal, Vol. 50, pp. 1969–2014, July-August 1971.

14. J. Makhoul, "Stable and Efficient Lattice Methods for Linear Prediction," IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-25, No. 5, pp. 423–428. October 1977.

15. F. Itakura and S. Saito, "Digital Filtering Techniques for Speech Analysis and Systems," Proceedings of the International Congress on Acoustics, Budapest, Hungary, paper 25-C-1, pp. 261–264, 1971

16. J. P. Burg, "Maximum Entropy Spectral Analysis," Ph.D. disseration, Geophysics Department, Stanford University, Stanford, CA, May 1975.

17. J. Makhoul, "Linear Prediction: A Tutorial Review," Proceedings IEEE, Vol. 63, pp. 561–580, April 1975.